On stability of discrete-time quantum filters

Pierre Rouchon

On stability of discrete-time quantum filters

Pierre Rouchon*

May 18, 2010

Fidelity is known to increase through a Kraus map: the fidelity between two density matrices is less than the fidelity between their images via a Kraus map. We prove here that, in average, the square of the fidelity is also increasing for a quantum filter: the square of the fidelity between the density matrix of the underlying Markov chain and the density matrix of its associated quantum filter is a super-martingale. Thus discrete-time quantum filters are stable processes and tend to forget their initial conditions.

1 Kraus maps and quantum Markov chains

Take the Hilbert space $S = \mathbb{C}^n$ of dimension $n > 0$ and consider a quantum channel described by the Kraus map (see [3], chapter 4)

$$\mathcal{K}(\rho) = \sum_{\mu=1}^{m} M_\mu \rho M_\mu^\dagger \tag{1}$$

where

- ρ is the density matrix describing the input quantum state, $\mathcal{K}(\rho)$ being then the output quantum state; $\rho \in \mathbb{C}^{n\times n}$ is a density matrix, i.e., an Hermitian matrix semi-positive definite and of trace one;
- for each $\mu \in \{1, \dots, m\}$, $M_\mu \in \mathbb{C}^{n\times n}/\{0\}$, and $\sum_\mu M_\mu^\dagger M_\mu = I$.

To this quantum channel is associated the following discrete-time Markov chain:

$$\rho_{k+1} = \mathcal{M}_{\mu_k}(\rho_k) \tag{2}$$

where

- ρ_k is the quantum state at sampling time t_k and k the sampling index ($t_k < t_{k+1}$).
- $\mu_k \in \{1, \dots, m\}$ is a random variable; $\mu_k = \mu$ with probability $p_\mu(\rho_k) = \mathrm{Tr}\left(M_\mu \rho_k M_\mu^\dagger\right)$.

*Mines ParisTech: pierre.rouchon@mines-paristech.fr

- $\mathcal{M}_\mu(\rho) = \frac{1}{\mathrm{Tr}\left(M_\mu \rho M_\mu^\dagger\right)} M_\mu \rho M_\mu^\dagger = \frac{1}{p_\mu(\rho)} M_\mu \rho M_\mu^\dagger$.

Kraus maps are contractions for the trace distance, i.e., nuclear distance (see [3], theorem 9.2, page 406): for all density matrices σ, ρ, one has

$$\mathrm{Tr}\left(|\mathcal{K}(\sigma) - \mathcal{K}(\rho)|\right) \leq \mathrm{Tr}\left(|\sigma - \rho|\right)$$

where, for any Hermitian matrix A with spectrum $\{\lambda_l\}_{l\in\{1,\dots,n\}}$, $\mathrm{Tr}\left(|A|\right) = \sum_{l=1}^{n} |\lambda_l|$. The Kraus map tends also to increase fidelity F (see [3], theorem 9.6, page 414): for all density matrices ρ and σ, one has

$$\mathrm{Tr}\left(\sqrt{\sqrt{\mathcal{K}(\sigma)}\mathcal{K}(\rho)\sqrt{\mathcal{K}(\sigma)}}\right) = F(\mathcal{K}(\sigma), \mathcal{K}(\rho)) \geq F(\sigma, \rho) = \mathrm{Tr}\left(\sqrt{\sqrt{\sigma}\rho\sqrt{\sigma}}\right) \tag{3}$$

where, for any Hermitian semi-positive matrix $A = U\Lambda U^\dagger$, U unitary matrix and $\Lambda = \mathrm{diag}\{\lambda_l\}_{l\in\{1,\dots,n\}}$, $\sqrt{A} = U\sqrt{\Lambda}U^\dagger$ with $\sqrt{\Lambda} = \mathrm{diag}\{\sqrt{\lambda_l}\}_{l\in\{1,\dots,n\}}$.

The conditional expectation of ρ_{k+1} knowing ρ_k is given by the Kraus map:

$$\mathbb{E}\left(\rho_{k+1}/\rho_k\right) = \mathcal{K}(\rho_k).$$

This result from the trivial identity $\sum_{\mu=1}^{m} \mathrm{Tr}\left(M_\mu \rho M_\mu^\dagger\right) \frac{M_\mu \rho M_\mu^\dagger}{\mathrm{Tr}\left(M_\mu \rho M_\mu^\dagger\right)}, = \mathcal{K}(\rho)$. In section 2, we show during the proof of theorem (1) the following inequality

$$\sum_{\mu=1}^{m} \mathrm{Tr}\left(M_\mu \rho M_\mu^\dagger\right) F^2\left(\frac{M_\mu \sigma M_\mu^\dagger}{\mathrm{Tr}\left(M_\mu \sigma M_\mu^\dagger\right)}, \frac{M_\mu \rho M_\mu^\dagger}{\mathrm{Tr}\left(M_\mu \rho M_\mu^\dagger\right)}\right) \geq F^2(\sigma, \rho) \tag{4}$$

for any density matrices ρ and σ. The left-hand side is related to a conditional expectation. Inequality (4), attached to the probabilistic mapping (2), can be seen as the stochastic counter-part of inequality (3) attached to the deterministic mapping (1). When for some μ, $\mathrm{Tr}\left(M_\mu \sigma M_\mu^\dagger\right) = 0$ with $\mathrm{Tr}\left(M_\mu \rho M_\mu^\dagger\right) > 0$, one term in the sum at the left-hand side of (4) is not defined. This is not problematic, since in this case, if we replace $\frac{M_\mu \sigma M_\mu^\dagger}{\mathrm{Tr}\left(M_\mu \sigma M_\mu^\dagger\right)}$ by $\frac{M_\mu \xi M_\mu^\dagger}{\mathrm{Tr}\left(M_\mu \xi M_\mu^\dagger\right)}$ where ξ is any density matrix such that $\mathrm{Tr}\left(M_\mu \xi M_\mu^\dagger\right) > 0$, this term is then well defined (in a multi-valued way) and inequality (4) remains satisfied for any such ξ.

During the proof of theorem (8), we extend this inequality to any partition of $\{1, \dots, m\}$ into $p \geq 1$ sub-sets $\mathcal{P}_\nu$:

$$\sum_{\nu=1}^{\nu=p} \mathrm{Tr}\left(\sum_{\mu\in\mathcal{P}_\nu} M_\mu \rho M_\mu^\dagger\right) F^2\left(\frac{\sum_{\mu\in\mathcal{P}_\nu} M_\mu \sigma M_\mu^\dagger}{\mathrm{Tr}\left(\sum_{\mu\in\mathcal{P}_\nu} M_\mu \sigma M_\mu^\dagger\right)}, \frac{\sum_{\mu\in\mathcal{P}_\nu} M_\mu \rho M_\mu^\dagger}{\mathrm{Tr}\left(\sum_{\mu\in\mathcal{P}_\nu} M_\mu \rho M_\mu^\dagger\right)}\right) \geq F^2(\sigma, \rho) \tag{5}$$

2 The standard case.

Take a realization of the Markov chain associated to the Kraus map $\mathcal{K}$. Assume that we detect, for each k, the jump μ_k but that we do not know the initial state ρ_0. The objective is to propose at sampling k, an estimation $\hat{\rho}_k$ of ρ_k based on the past detections $\mu_0, \ldots, \mu_{k-1}$. The simplest method consists in starting from an initial estimation $\hat{\rho}_0$ and at each sampling step to jump according to the detection. This leads to the following estimation scheme known as *a quantum filter*:

$$\hat{\rho}_{k+1} = \mathcal{M}_{\mu_k}(\hat{\rho}_k) \tag{6}$$

with $p_\mu(\rho_k) = \mathrm{Tr}\left(M_\mu \rho_k M_\mu\right)$ as probability of $\mu_k = \mu$. Notice that when $\mathrm{Tr}\left(M_{\mu_k} \hat{\rho}_k M_{\mu_k}\right) = 0$, $\mathcal{M}_{\mu_k}(\hat{\rho}_k)$ is not defined and should be replaced by $\mathcal{M}_{\mu_k}(\xi)$ where ξ is any density matrix such that $\mathrm{Tr}\left(M_{\mu_k} \hat{\xi} M_{\mu_k}\right) > 0$ (take, e.g., $\xi = \frac{1}{n} I_d$). The theorem here below is a first step to investigate the convergence of $\hat{\rho}_k$ towards ρ_k as k increases.

Theorem 1. *Consider the Markov chain of state* $(\rho_k, \hat{\rho}_k)$ *satisfying* (2) *and* (6). *Then* $F^2(\hat{\rho}_k, \rho_k)$ *is a super-martingale:* $\mathbb{E}\left(F^2(\hat{\rho}_{k+1}, \rho_{k+1}) / (\hat{\rho}_k, \rho_k)\right) \geq F^2(\hat{\rho}_k, \rho_k)$.

When $\hat{\rho}_k$ or ρ_k are pure states, $\hat{\rho}_{k+1}$ or ρ_{k+1} remain also a pure states. Then, $F^2(\hat{\rho}_k, \rho_k) = \mathrm{Tr}\left(\hat{\rho}_k \rho_k\right)$ and $F^2(\hat{\rho}_{k+1}, \rho_{k+1}) = \mathrm{Tr}\left(\hat{\rho}_{k+1}, \rho_{k+1}\right)$. In this case, theorem 1 has been proved in [2] using Cauchy-Schwartz inequalities for $m = 2$. The proof proposed here below deals with the general case when both ρ_k and $\hat{\rho}_k$ can be mixed states. It relies on arguments similar to those used for the proof of theorem 9.6 in [3].

Proof. ρ and $\hat{\rho}$ are associated to the Hilbert space $S = \mathbb{C}^n$: ρ and $\hat{\rho}$ are operators from S to S. Take a copy $Q = \mathbb{C}^n$ of S and consider the composite system living on $S \otimes Q \equiv \mathbb{C}^{n^2}$. Then $\hat{\rho}$ and ρ correspond to partial traces versus Q of projectors $|\hat{\psi}\rangle\langle\hat{\psi}|$ and $|\psi\rangle\langle\psi|$ associated to pure states $|\hat{\psi}\rangle$ and $|\psi\rangle \in S \otimes Q$:

$$\hat{\rho} = \mathrm{Tr}_Q\left(|\hat{\psi}\rangle\langle\hat{\psi}|\right), \quad \rho = \mathrm{Tr}_Q\left(|\psi\rangle\langle\psi|\right)$$

$|\hat{\psi}\rangle$ and $|\psi\rangle$ are called purifications of $\hat{\rho}$ and ρ. They are not unique but one can always choose them such that $F(\hat{\rho}, \rho) = |\langle\hat{\psi}|\psi\rangle|$ (Uhlmann's theorem).

Denote by $|\hat{\psi}_k\rangle$ and $|\psi_k\rangle$ such purifications of $\hat{\rho}_k$ and ρ_k satisfying $F(\hat{\rho}_k, \rho_k) = |\langle\hat{\psi}_k|\psi_k\rangle|$. We have

$$\mathbb{E}\left(F^2(\hat{\rho}_{k+1}, \rho_{k+1}) / (\hat{\rho}_k, \rho_k)\right) = \sum_{\mu=1}^{m} p_\mu(\rho_k) F^2(\mathcal{M}_\mu(\hat{\rho}_k), \mathcal{M}_\mu(\rho_k)).$$

The matrices $\mathcal{M}_\mu(\hat{\rho}_k)$ and $\mathcal{M}_\mu(\rho_k)$ are also density matrices. Take the space $S \otimes Q \otimes E$ where E is the Hilbert space of the environment appearing in the system-environment model of the Kraus map (1). This model is recalled in appendix A. It introduced an unitary transformation U on $S \otimes E$. This unitary transformation can be extended to $S \otimes Q \otimes E \equiv S \otimes E \otimes Q$ by setting $V = U \otimes I$ (I is identity on Q). Then

$$M_\mu(\rho_k) = \mathrm{Tr}_{Q \otimes E}\left(P_\mu V \left(|\psi_k\rangle\langle\psi_k| \otimes |e_0\rangle\langle e_0|\right) V^\dagger P_\mu\right).$$

Set $|\phi_k\rangle = |\psi_k\rangle \otimes |e_0\rangle \in S \otimes Q \otimes E$ and $|\chi_k\rangle = V|\phi_k\rangle$. Using $P_\mu^2 = P_\mu$, we have

$$p_\mu(\rho_k) = \mathrm{Tr}\left(M_\mu(\rho_k)\right) = \langle \phi_k | V^\dagger P_\mu V | \phi_k \rangle = \| P_\mu |\chi_k\rangle \|^2$$

For each μ, the state $|\chi_{k\mu}\rangle = \frac{1}{\sqrt{p_\mu(\rho_k)}} P_\mu |\chi_k\rangle$ is a purification of $\mathcal{M}_\mu(\rho_k)$:

$$\mathcal{M}_\mu(\rho_k) = \mathrm{Tr}_{Q\otimes E}\left(|\chi_{k\mu}\rangle\langle\chi_{k\mu}|\right).$$

Similarly set $|\hat{\phi}_k\rangle = |\hat{\psi}_k\rangle \otimes |e_0\rangle$ and $|\hat{\chi}_k\rangle = V|\hat{\phi}_k\rangle$. For each μ, $|\hat{\chi}_{k\mu}\rangle = \frac{1}{\sqrt{p_\mu(\hat{\rho}_k)}} P_\mu |\hat{\chi}_k\rangle$ is also a purification of $\mathcal{M}_\mu(\hat{\rho}_k)$. By Uhlmann's theorem,

$$F^2(\mathcal{M}_\mu(\hat{\rho}_k), \mathcal{M}_\mu(\rho_k)) \geq |\langle \hat{\chi}_{k\mu} | \chi_{k\mu} \rangle|^2.$$

Thus we have

$$\mathbb{E}\left(F^2(\hat{\rho}_{k+1}, \rho_{k+1}) / (\hat{\rho}_k, \rho_k)\right) \geq \sum_{\mu=1}^{m} p_\mu(\rho_k)\ |\langle \hat{\chi}_{k\mu} | \chi_{k\mu} \rangle|^2.$$

Since V is unitary,

$$|\langle \hat{\chi}_k | \chi_k \rangle|^2 = |\langle \hat{\phi}_k | \phi_k \rangle|^2 = |\langle \hat{\psi}_k | \psi_k \rangle|^2 = F^2(\hat{\rho}_k, \rho_k).$$

Let us show that $\sum_{\mu=1}^m p_\mu(\rho_k)\ |\langle \hat{\chi}_{k\mu} | \chi_{k\mu} \rangle|^2 \geq |\langle \hat{\chi}_k | \chi_k \rangle|^2$. We have

$$p_\mu(\rho_k)\ |\langle \hat{\chi}_{k\mu} | \chi_{k\mu} \rangle|^2 = |\langle \hat{\chi}_{k\mu} | P_\mu \chi_k \rangle|^2 = |\langle \hat{\chi}_{k\mu} | \chi_k \rangle|^2,$$

thus it is enough to prove that $\sum_{\mu=1}^m |\langle \hat{\chi}_{k\mu} | \chi_k \rangle|^2 \geq |\langle \hat{\chi}_k | \chi_k \rangle|^2$. Denote by $\hat{R} \subset S \otimes Q \otimes E$ the vector space spanned by the ortho-normal basis $(|\hat{\chi}_{k\mu}\rangle)_{\mu\in\{1,\ldots,m\}}$ and by $\hat{P}$ the projector on $\hat{R}$. Since

$$|\hat{\chi}_k\rangle = \sum_{\mu=1}^m P_\mu |\hat{\chi}_k\rangle = \sum_{\mu=1}^m \sqrt{p_\mu(\hat{\rho}_k)} |\hat{\chi}_{k\mu}\rangle$$

$|\hat{\chi}_k\rangle$ belongs to $\hat{R}$ and thus $|\langle \hat{\chi}_k | \chi_k \rangle|^2 = |\langle \hat{\chi}_k | \hat{P} | \chi_k \rangle\rangle|^2$. We conclude by Cauchy-Schwartz inequality

$$|\langle \hat{\chi}_k | \chi_k \rangle|^2 = |\langle \hat{\chi}_k | \hat{P} | \chi_k \rangle\rangle|^2 \leq \|\hat{\chi}_k\|^2 \|\hat{P} |\chi_k\rangle\|^2 = \|\hat{P} |\chi_k\rangle\|^2 = \sum_{\mu=1}^m |\langle \hat{\chi}_{k\mu} | \chi_k \rangle|^2.$$

□

3 The aggregated case.

Let us consider another Markov chain attached to the same Kraus map (1) and associated to a partition of $\{1, \ldots, m\}$ into $p \geq 1$ sub-sets $\mathcal{P}_\nu$ (aggregation of several quantum jumps via "partial Kraus maps"):

$$\rho_{k+1} = \frac{1}{\mathrm{Tr}\left(\sum_{\mu\in\mathcal{P}_{\nu_k}} M_\mu \rho_k M_\mu^\dagger\right)} \left(\sum_{\mu\in\mathcal{P}_{\nu_k}} M_\mu \rho_k M_\mu^\dagger\right) \tag{7}$$

where $\nu_k = \nu$ with probability $\mathrm{Tr}\left(\sum_{\mu\in\mathcal{P}_\nu} M_\mu \rho_k M_\mu^\dagger\right)$. Consider the associated quantum filter

$$\hat{\rho}_{k+1} = \frac{1}{\mathrm{Tr}\left(\sum_{\mu\in\mathcal{P}_{\nu_k}} M_\mu \hat{\rho}_k M_\mu^\dagger\right)} \left(\sum_{\mu\in\mathcal{P}_{\nu_k}} M_\mu \hat{\rho}_k M_\mu^\dagger\right) \tag{8}$$

where the jump index ν_k coincides with the jump index ν_k in (7). Then we have the following theorem.

Theorem 2. *Consider the Markov chain of state $(\rho_k, \hat{\rho}_k)$ satisfying (7) and (8). Then $F^2(\hat{\rho}_k, \rho_k)$ is a super-martingale: $\mathbb{E}\left(F^2(\hat{\rho}_{k+1}, \rho_{k+1})/(\hat{\rho}_k, \rho_k)\right) \geq F^2(\hat{\rho}_k, \rho_k)$.*

Proof. It is similar to the proof of theorem 1. We will just point out here the main changes using the same notations. We start from

$$\mathbb{E}\left(F^2(\hat{\rho}_{k+1}, \rho_{k+1})/(\hat{\rho}_k, \rho_k)\right) = \sum_{\nu=1}^{p} \tilde{p}_\nu(\rho_k) F^2(\tilde{\mathcal{M}}_\nu(\hat{\rho}_k), \tilde{\mathcal{M}}_\nu(\rho_k)).$$

where we have set

$$\tilde{p}_\nu(\rho) = \mathrm{Tr}\left(\sum_{\mu\in\mathcal{P}_\nu} M_\mu \rho M_\mu^\dagger\right), \quad \tilde{\mathcal{M}}_\nu(\rho) = \tfrac{1}{\tilde{p}_\nu(\rho)} \left(\sum_{\mu\in\mathcal{P}_\nu} M_\mu \rho M_\mu^\dagger\right).$$

With $\tilde{P}_\nu$ the orthogonal projector on $S\otimes Q\otimes \mathrm{span}\{|\mu\rangle, \mu \in \mathcal{P}_\nu\}$ and $\tilde{M}_\nu(\rho) = \sum_{\mu\in\mathcal{P}_\nu} M_\mu \rho M_\mu^\dagger$, we have

$$\tilde{M}_\nu(\rho_k) = \mathrm{Tr}_{Q\otimes E}\left(\tilde{P}_\nu V\left(|\psi_k\rangle\langle\psi_k| \otimes |e_0\rangle\langle e_0|\right) V^\dagger \tilde{P}_\nu\right)$$

and

$$\tilde{p}_\nu(\rho_k) = \mathrm{Tr}\left(\tilde{M}_\nu(\rho_k)\right) = \langle\phi_k|V^\dagger \tilde{P}_\nu V|\phi_k\rangle = \|\tilde{P}_\nu|\chi_k\rangle\|^2$$

For each ν, the state $|\tilde{\chi}_{k\nu}\rangle = \frac{1}{\sqrt{\tilde{p}_\nu(\rho_k)}}\tilde{P}_\nu|\chi_k\rangle$ is a purification of $\tilde{\mathcal{M}}_\nu(\rho_k)$:

$$\tilde{\mathcal{M}}_\nu(\rho_k) = \mathrm{Tr}_{Q\otimes E}\left(|\tilde{\chi}_{k\nu}\rangle\langle\tilde{\chi}_{k\nu}|\right).$$

Similarly $|\hat{\tilde{\chi}}_{k\nu}\rangle = \frac{1}{\sqrt{\tilde{p}_\nu(\hat{\rho}_k)}}\tilde{P}_\nu|\hat{\chi}_k\rangle$ is also a purification of $\tilde{\mathcal{M}}_\nu(\hat{\rho}_k)$. By Uhlmann's theorem,

$$F^2(\tilde{\mathcal{M}}_\nu(\hat{\rho}_k), \tilde{\mathcal{M}}_\nu(\rho_k)) \geq |\langle\hat{\tilde{\chi}}_{k\nu}|\tilde{\chi}_{k\nu}\rangle|^2.$$

Thus we have

$$\mathbb{E}\left(F^2(\hat{\rho}_{k+1}, \rho_{k+1})/(\hat{\rho}_k, \rho_k)\right) \geq \sum_{\nu=1}^{p} \tilde{p}_\nu(\rho_k)\ |\langle\hat{\tilde{\chi}}_{k\mu}|\tilde{\chi}_{k\mu}\rangle|^2.$$

Let us show that $\sum_{\nu=1}^{p} \tilde{p}_\nu(\rho_k)\ |\langle\hat{\tilde{\chi}}_{k\nu}|\tilde{\chi}_{k\nu}\rangle|^2 \geq |\langle\hat{\chi}_k|\chi_k\rangle|^2 = F^2(\hat{\rho}_k, \rho_k)$. We have

$$\tilde{p}_\nu(\rho_k)\ |\langle\hat{\tilde{\chi}}_{k\nu}|\tilde{\chi}_{k\nu}\rangle|^2 = |\langle\hat{\tilde{\chi}}_{k\nu}|\tilde{P}_\nu\chi_k\rangle|^2 = |\langle\hat{\tilde{\chi}}_{k\nu}|\chi_k\rangle|^2,$$

thus it is enough to prove that $\sum_{\nu=1}^{p} |\langle \hat{\tilde{\chi}}_{k\nu}|\chi_k\rangle|^2 \geq |\langle \hat{\chi}_k|\chi_k\rangle|^2$. Denote by $\hat{\tilde{R}} \subset S \otimes Q \otimes E$ the vector space spanned by the ortho-normal basis $\left(|\hat{\tilde{\chi}}_{k\nu}\rangle\right)_{\nu \in \{1,\ldots,p\}}$ and by $\hat{\tilde{P}}$ the projector on $\hat{\tilde{R}}$. Since

$$|\hat{\chi}_k\rangle = \sum_{\nu=1}^{p} \tilde{P}_\nu |\hat{\chi}_k\rangle = \sum_{\nu=1}^{p} \sqrt{\tilde{p}_\nu(\hat{\rho}_k)}|\hat{\tilde{\chi}}_{k\nu}\rangle$$

$|\hat{\chi}_k\rangle$ belongs to $\hat{\tilde{R}}$ and thus $|\langle \hat{\chi}_k|\chi_k\rangle|^2 = |\langle \hat{\chi}_k|\hat{\tilde{P}}|\chi_k\rangle\rangle|^2$. We conclude by Cauchy-Schwartz inequality

$$|\langle \hat{\chi}_k|\chi_k\rangle|^2 = |\langle \hat{\chi}_k|\hat{\tilde{P}}|\chi_k\rangle\rangle|^2 \leq \|\hat{\chi}_k\|^2 \|\hat{\tilde{P}}|\chi_k\rangle\|^2 = \|\hat{\tilde{P}}|\chi_k\rangle\|^2 = \sum_{\nu=1}^{p} |\langle \hat{\tilde{\chi}}_{k\nu}|\chi_k\rangle|^2.$$

$\square$

4 Concluding remarks

Theorems 1 and 2 are still valid if the Kraus operators M_μ depend on k. In particular, $F^(\hat{\rho}_k, \rho_k)$ remains a super-martingale even if the Kraus operators depend on $\hat{\rho}_k$, i.e., in case of feedback.

When σ and ρ are pure states (projectors of rank one), $D(\sigma,\rho) = \sqrt{1 - F^2(\sigma,\rho)}$. Consequently inequality (4) yields to

$$\sum_{\mu=1}^{m} \mathrm{Tr}\left(M_\mu \rho M_\mu^\dagger\right) D\left(\frac{M_\mu \sigma M_\mu^\dagger}{\mathrm{Tr}\left(M_\mu \sigma M_\mu^\dagger\right)}, \frac{M_\mu \rho M_\mu^\dagger}{\mathrm{Tr}\left(M_\mu \rho M_\mu^\dagger\right)}\right) \leq D(\sigma,\rho)$$

for any pure states σ and ρ (use the fact that $[0,x] \ni x \mapsto \sqrt{1-x}$ is decreasing and concave). We conjecture that such inequality hold also true for any mixed states and that $D(\hat{\rho}_k, \rho_k) = \mathrm{Tr}\left(|\hat{\rho}_k - \rho_k|\right)$ is a sub-martingale.

References

[1] S. Haroche and J.M. Raimond. *Exploring the Quantum: Atoms, Cavities and Photons.* Oxford University Press, 2006.

[2] M. Mirrahimi, I. Dotsenko, and P. Rouchon. Feedback generation of quantum Fock states by discrete QND measures. *CDC 2009, Shanghai*, 2009.

[3] M.A. Nielsen and I.L. Chuang. *Quantum Computation and Quantum Information.* Cambridge University Press, 2000.

A System-environment model

The quantum channel associated to the Kraus map (1) or the Markov chain (2) admits a system-environment model (see [1], chapter 4 entitled "The environment is watching"). Take the Hilbert space $E = \mathbb{C}^m$ associated to the environment and the composite system living on $S \otimes E$. Take a pure state $|\phi_k\rangle \in S$ and its density matrix $\rho_k = |\phi_k\rangle\langle\phi_k|$. Assume that before detection μ_k at step k, the composite system admits the pure state $|\phi_k\rangle \otimes |e_0\rangle$ where $|e_0\rangle$ is an environment pure state. Take m states $|\mu\rangle$ forming an orthogonal base of E. Then exists a unitary transformation U (not unique) of $S \otimes E$ such that, for all $|\phi\rangle \in S$,

$$U\left(|\phi\rangle \otimes |e_0\rangle\right) = \sum_{\mu=1}^{m} \left(M_\mu|\phi\rangle\right) \otimes |\mu\rangle.$$

This is a direct consequence of $\sum_{\mu=1}^{n} M_\mu^\dagger M_\mu = I$. For each μ, denote by P_μ the orthogonal projector onto the subspace $S \otimes (\mathbb{C}|\mu\rangle)$. Then $P_\mu U\left(|\phi\rangle \otimes |e_0\rangle\right) = \left(M_\mu|\phi\rangle\right) \otimes |\mu\rangle$ and $\sum_\mu P_\mu = I$. We can then verify that for any density matrix ρ associated to a state in R,

$$P_\mu U\left(\rho \otimes |e_0\rangle\langle e_0|\right) U^\dagger P_\mu = M_\mu \rho M_\mu^\dagger \otimes |\mu\rangle\langle\mu|$$

and thus

$$\mathrm{Tr}_E\left(P_\mu U\left(\rho \otimes |e_0\rangle\langle e_0|\right) U^\dagger P_\mu\right) = M_\mu \rho M_\mu^\dagger.$$